

# Northumbria Research Link

Citation: Wu, Ruiqi, Zhou, Changle, Chao, Fei, Zhu, Zuyuan, Lin, Chih-Min and Yang, Longzhi (2017) A Developmental Learning Approach of Mobile Manipulator via Playing. Frontiers in Neurorobotics, 11. p. 53. ISSN 1662-5218

Published by: Frontiers

URL: <https://doi.org/10.3389/fnbot.2017.00053>  
<<https://doi.org/10.3389/fnbot.2017.00053>>

This version was downloaded from Northumbria Research Link:  
<http://nrl.northumbria.ac.uk/id/eprint/32386/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)



**Northumbria  
University**  
NEWCASTLE



**UniversityLibrary**



# A Developmental Learning Approach of Mobile Manipulator via Playing

Ruiqi Wu<sup>1</sup>, Changle Zhou<sup>1</sup>, Fei Chao<sup>1\*</sup>, Zuyuan Zhu<sup>2</sup>, Chih-Min Lin<sup>1,3</sup> and Longzhi Yang<sup>4</sup>

<sup>1</sup> Fujian Provincial Key Lab of Brain-Inspired Computing, Department of Cognitive Science, School of Informatics, Xiamen University, Xiamen, China, <sup>2</sup> Department of Computer Science, School of Computer Science and Electronic Engineering, University of Essex, Colchester, United Kingdom, <sup>3</sup> Department of Electrical Engineering, Yuan Ze University, Tao-Yuan, Taiwan, <sup>4</sup> Department of Computer and Information Sciences, Faculty of Engineering and Environment, Northumbria University, Newcastle upon Tyne, United Kingdom

Inspired by infant development theories, a robotic developmental model combined with game elements is proposed in this paper. This model does not require the definition of specific developmental goals for the robot, but the developmental goals are implied in the goals of a series of game tasks. The games are characterized into a sequence of game modes based on the complexity of the game tasks from simple to complex, and the task complexity is determined by the applications of developmental constraints. Given a current mode, the robot switches to play in a more complicated game mode when it cannot find any new salient stimuli in the current mode. By doing so, the robot gradually achieves its developmental goals by playing different modes of games. In the experiment, the game was instantiated into a mobile robot with the playing task of picking up toys, and the game is designed with a simple game mode and a complex game mode. A developmental algorithm, “Lift-Constraint, Act and Saturate,” is employed to drive the mobile robot move from the simple mode to the complex one. The experimental results show that the mobile manipulator is able to successfully learn the mobile grasping ability after playing simple and complex games, which is promising in developing robotic abilities to solve complex tasks using games.

**Keywords:** developmental robotics, mobile manipulator, robotic hand-eye coordination, neural network control, sensory-motor coordination

## OPEN ACCESS

### Edited by:

Patricia Shaw,  
Aberystwyth University,  
United Kingdom

### Reviewed by:

Eiji Uchibe,  
Advanced Telecommunications  
Research Institute International, Japan  
Daniele Caligiore,  
Consiglio Nazionale Delle Ricerche  
(CNR), Italy

### \*Correspondence:

Fei Chao  
fchao@xmu.edu.cn

**Received:** 05 June 2017

**Accepted:** 19 September 2017

**Published:** 04 October 2017

### Citation:

Wu R, Zhou C, Chao F, Zhu Z,  
Lin C-M and Yang L (2017) A  
Developmental Learning Approach of  
Mobile Manipulator via Playing.  
Front. Neurorobot. 11:53.  
doi: 10.3389/fnbot.2017.00053

## 1. INTRODUCTION

Intelligent robots have been widely applied to support or even replace the work of humans in many social activities, such as assembly lines, family services, and social entertainment. These robots are made intelligent by many methods proposed in the literature, with the most common ones being mathematical modeling and dynamics models, such as Yan et al. (2013), Galbraith et al. (2015) and Grinke et al. (2015). These methods utilize predefined cognitive architectures in the intelligent systems, which cannot be used for significant changes during the interaction within the environment. If the intelligent system is applied in a new environment, the intelligent systems must be reconstructed. Also the complexity of the model increases exponentially as the complexity of the task increases. In addition, it is still very challenging in the field of robotics to allow the robot to learn complex skills and incorporate a variety of skills in an intelligent system.

Asada et al. (2001), Lungarella et al. (2003), and Weng (2004) attempt to let the robot learn intricate skills using the so-called developmental robotics approaches. These approaches enable

robots to gradually develop multiple basic skills and thus learn to handle complex tasks (Berthouze and Lungarella, 2004; Jiang et al., 2014). In other words, the learning target of a complex set of skills are divided into the learning of a number of stage targets (Wang et al., 2013; Zhu et al., 2015), and the robot achieves the ultimate learning goal by completing a series of sub learning goals. This method reduces the difficulty for the robot to learn new skills (Shaw et al., 2014), and gives the robot the ability to accumulate learning, where the basic skills learned during the development process are reserved so as to arrive at the final skill (Lee et al., 2013). When a robot uses the method of developmental robotics to learn new skills, the target in every developmental phase must be clearly defined (Stoytchev, 2009). However, this is practically very challenge for those phases with a large number of complex tasks, thereby limiting the applicability of developmental robotics.

It has been observed by infant development researchers that infants and young children, when developing skills, do not need to define specific developmental goals (Adolph and Joh, 2007), and mergence is the primary form for infants to acquire skills (Morse and Cangelosi, 2017). In particular, a play phenomenon often accompanies the process of an infant skill development (Cangelosi et al., 2015), which has led to one infant development theory that infants develop relevant skills during play. The play of the early infant is driven primarily by intrinsic motivation (Oudeyer et al., 2007; Baldassarre and Mirolli, 2013; Caligiore et al., 2015), and an infant's development goal is implied in the game that the infant plays. This theory has not been applied and verified in developmental robotics. Therefore, a robotic developmental model that combines the infant developmental theory and developmental robotics is proposed herein. In this model, the learning skills of a robot are artificially viewed as game playing by an infant, and the developmental target is implied in the game goals. Then, a method of developmental robotics is utilised by the model to accomplish the robot's skill development and learning. The proposed system not only reduces the difficulty of robot learning and allows accumulate learning, but also mitigates the limitation of applicability as discussed above by clearly defining goals in the developmental method.

In contrast to other developmental learning methods (Yang and Asada, 1996; Berthouze and Lungarella, 2004), the proposed approach embeds the role of play in early infant development into the developmental learning approach. Through two game modes, our robot developed mobile reaching and grasping abilities with no external reward existing in the two game modes. The robot merely uses its learning status to switch from one game mode into next one. Such approach also adopts the intrinsic motivation-driven learning method. Therefore, the main contribution of this work is a combined developmental algorithm that allows robot to acquire new abilities by applying the infant developmental theory, in which skills are developed through playing. With the inclusion of game elements, the robot can acquire developed mobile reaching and grasping skills with emergence.

The remainder of this paper is organized as follows: section 2 introduces the background knowledge of developmental robotics and the "Lift-Constraint, Act and Saturate" developmental algorithm. Section 3 outlines our model and designs the

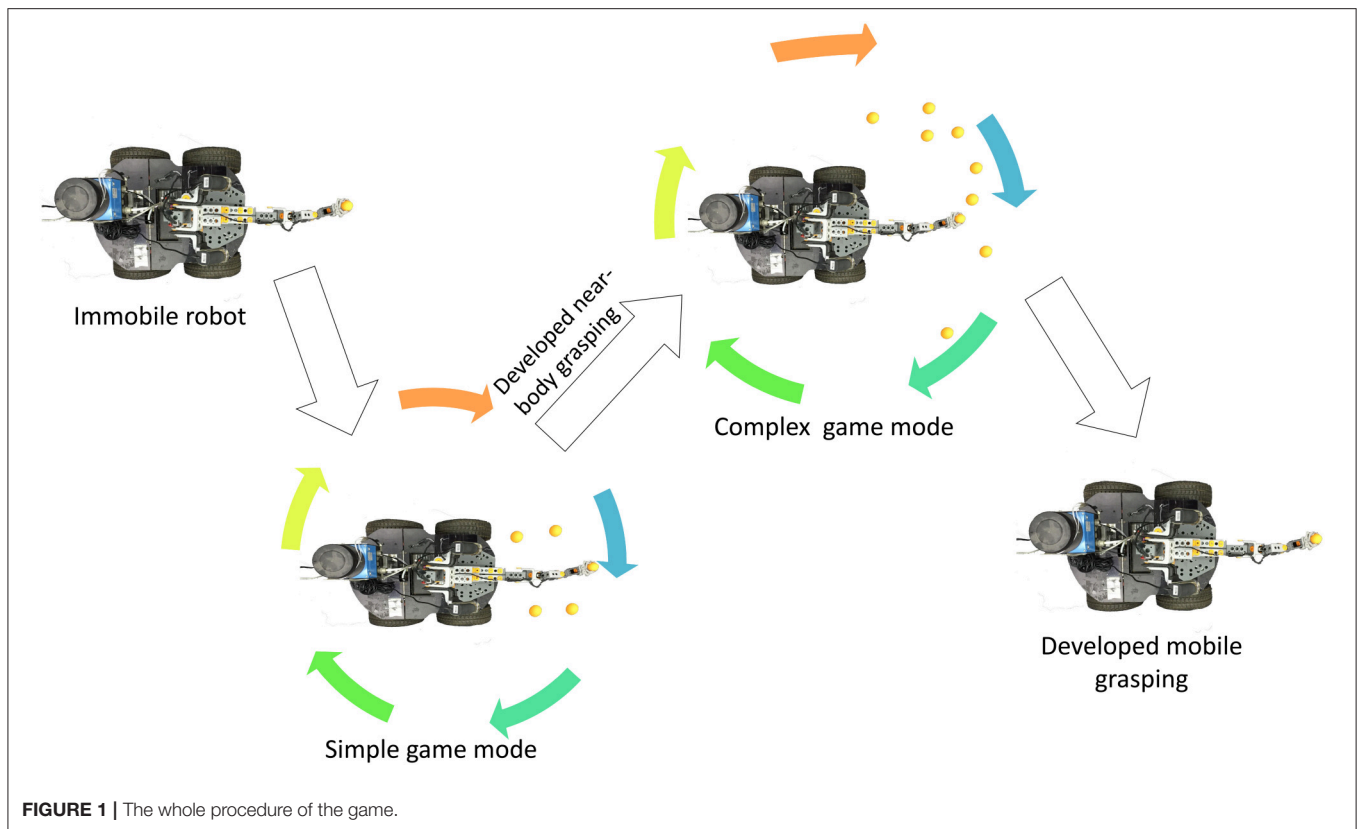
developmental strategy of robots. Section 4 describes the experimentation and analyzes the results. Section 5 concludes the paper and points out possible future work.

## 2. DEVELOPMENTAL ROBOTICS

As a research method with an interdisciplinary background of developmental psychology, neuroscience, computer science, etc. (Earland et al., 2014; Law et al., 2014a; Gogate, 2016), developmental robotics aims to provide solutions in the design of behavior and cognition in the artificial intelligence systems (Marocco et al., 2010; Baillie, 2016; Salgado et al., 2016). Developmental robotics is inspired by the developmental principles and mechanisms observed during the development of infants and children (Chao et al., 2014a), and thus the main idea of developmental robotics is to let a robot imitate a human's development process (Adolph and Joh, 2007; Oudeyer, 2017). The robot achieves sensory-movement and cognitive ability of incremental acquisition according to the inherent development principles and through real-time interaction with the external environment (Cangelosi et al., 2015). Developmental robotics focuses on two primary challenges in the field of robotics: (1) learning new knowledge and skills from a constantly changing environment; and (2) understanding their relationship with their physical environment and other agents.

Guerin et al. (2013) suggested in developmental robotics that most patterns need to be learned from a few patterns and the described knowledge must be developed gradually, by alluding to the general mechanism of sensory-movement development and the knowledge description in action-object relationships. Law et al. (2014a) achieved stage development on an iCub robot. They successfully built a development model for infants from birth to 6 months, which is driven by a new control system. Starting from uncontrolled movements and passing through several obvious stages of behavior, the iCub robot, like an infant, finally reaches out and simply manipulates the object. Cangelosi et al. (2015) used a method of action-centering to perform a large number of synchronous comparisons with similar human development and artificial systems. They discovered that human development and artificial developmental systems share some common practices from which they can learn. These studies inspired the establishment of the proposed robotic systems reported in this paper using the key features and important theories in human infant development.

One of the two most important research focuses in the field of developmental robotic is the development of skills corresponding to a particular stage of an infant's development (Chao et al., 2010; Law et al., 2013), and another is the modeling of the multi-stage development process (Hülse et al., 2010; Law et al., 2014b). However, it is also of significant importance to study the impact of play in early infant development, which may also provide solutions in developmental robotics. Hart and Grupen (2011) proposed a robot which organizes its own sensory-movement space for incremental growth. Their solution uses internal incentive motivations to allow robots to assimilate new skills which are learned from the new learning phase or the new



environment and become the basis for the next learning phase. This research has been applied to humanoid robots. Benefiting from important theories on developmental psychology, those humanoid platforms can easily reproduce several behavioral patterns or validate new hypotheses. Savastano and Nolfi (2013) used a neural robot to imitate an infant's learning process, which was demonstrated by a humanoid robot incrementally learning the ability to grasp. In the experiment, the maturity limit was systematically controlled and a variety of developmental strategies were produced. The experiment also shows that human and robots can learn from each other by a comparative study. Different with these studies, the experiment platform presented in this work is a wheeled mobile robot, aiming to study the influence of play in developmental learning methods.

The "Lift-Constraint, Act and Saturate" (LCAS) approach (Lee et al., 2007), as a developmental learning algorithm, has been widely applied in the developmental robot system (Chao et al., 2013; Wang et al., 2014). The LCAS approach contains a loop with three segments: (1) Lift-Constraint, (2) Act, and (3) Saturate. First, all possible (or available) restrictions are stated clearly and their release times are formulated. Then, the robot learns that all existing constraints are substantiated. When the saturation rate of a robot's learning system is stable, a new constraint is released. From this, the robot learns either new knowledge or skills leading to the removal of a new environmental constraint. When the robot has lifted all constraints through learning, and all saturation rates of the robot learning system are stable, the robot has successfully learned a series of skills.

### 3. THE PROPOSED METHOD

#### 3.1. Model Overview

A developmental algorithm is proposed in this work by designing a game for robots that allows robots to develop skills by playing. Infants' stable grasping abilities are developed through grasping objects around them and infants are not happy until they can stably do so. Due to the constraints of body, infants pay most of their attentions on the range of physical activities while learning skills, which leads to more efficient skill learning. In the process of learning, infants do not have a clear learning goal, and all the activities are simply driven by intrinsic motivation and interest. Inspired by this, in our model, we design a game of pick-up-toys for a mobile manipulator, in which victory is the mobile manipulator successfully picking up surrounding toys. By playing this game, a robot with a mobile manipulator gradually develops mobile grasping ability. The game has two modes based on task complexity:

- **Simple game mode:** All toys are distributed within the working range of the manipulator, and the robot can perform toy grasping without moving.
- **Complex game mode:** Some toys are distributed beyond the working range of the manipulator, and the robot must move in order to grasp these toy.

The robot's skill development process is illustrated in **Figure 1**. The robot with the mobile manipulator is initialised to play in the simple game mode until it successfully completes the game. After

**TABLE 1 |** Infant's developmental stages and the corresponding development abilities.

The developmental stage	Developed ability
Visual fixation period	Fixation, saccade
Hand-eye coordination period	Hand-eye coordination, near-body grasping
Mobile-direction coordination period	Mobile grasping

the robot acquires near-body grasping ability, the game switches to the complex mode. The game is over when the robot acquires mobile grasping ability in this game mode.

The key in implementing a mobile robot with grasping ability is the coordination of the visual system, the manipulator, and the mobile system. This is mainly implemented by two basic non-linear mappings: (1) from the robot's visual space to the manipulator's movement space, and (2) from the visual space to the robot's mobile space. The two basic mappings are simulated in this work by two Radial Basis Function networks (RBF), and the training of these neural networks is accomplished by the robot playing the game.

### 3.2. Developmental Strategy

Before acquiring mobile grasping ability, infants must attain a number of developmental stages as listed in **Table 1** (Law et al., 2010, 2011), in which they develop a variety of basic abilities. A developmental strategy is designed to support the development of a robot's mobile grasping ability, which is implemented by the LCAS algorithm (i.e., Algorithm 1 shown below) (Chao et al., 2014b). In this pseudo code,  $i$  is the number of learning epochs under current constraints, and  $Sat(i)$  is the saturation rate in the  $i^{th}$  learning epoch. If the  $Sat(i)$  is true, the algorithm ends the training under the current constraint, and releases a new constraint. The value of  $Sat(i)$  is determined by Equation 1, where  $i$  is the number of training epochs;  $G(i)$  is the model's global excitation value at epoch  $i$ ;  $\epsilon$  controls the sampling rate; and the  $\phi$  is a fixed value used to control the global excitation's amplitude of variation. If the value of  $G(i)$  is  $< \phi$  and the variation of the global excitation is  $< \epsilon$ , a new constraint is lifted. In this work, the values of parameters  $\psi$ ,  $\epsilon$  and  $\phi$  are empirically set to 10, 0.5, and 0.02, respectively, and theoretical study on these parameters remain as future work.

$$Sat(i) = \begin{cases} true; & \text{if } |G(i) - G(i - \psi)| < \epsilon \\ & \text{and } G(i) < \phi; i = \psi \cdots n \\ false; & \text{else} \end{cases} \quad (1)$$

In the LCAS algorithm, a robot's constraints are first substantiated, as shown in **Table 2**. Once the constraints are substantiated, the development of the robot proceeds according to the lift-constraint strategy as listed in **Table 3**.

The LCAS algorithm in this work is executed in the following five steps: (1) The target object is placed in the robot's external environment. The robot acquires image information about the environment by removing the "visual resolution" constraint of the robot's eyes. (2) After the robot attains watch ability, the

#### Algorithm 1 Combined Learning Algorithm

```

1: while not all constraints are released do
2:   for  $i = 0$  to  $n$  do
3:     if  $Sat(i)$  is TRUE then
4:       quit this for-loop and release a new constraint;
5:     else
6:       repeat doing the leaning process within this for-loop
         (Lines 2-8);
7:     end if
8:   end for
9: end while

```

**TABLE 2 |** The constraint instantiation for mobile robot.

Constraint type	Substantiation
Hardware	Eye joint, arm joint, wheel motor
Sensory-motor	Tactile sensor, arm movement range
Cognitive	Neural network
Maturational	Network convergence threshold
External/environmental	Location of the target object

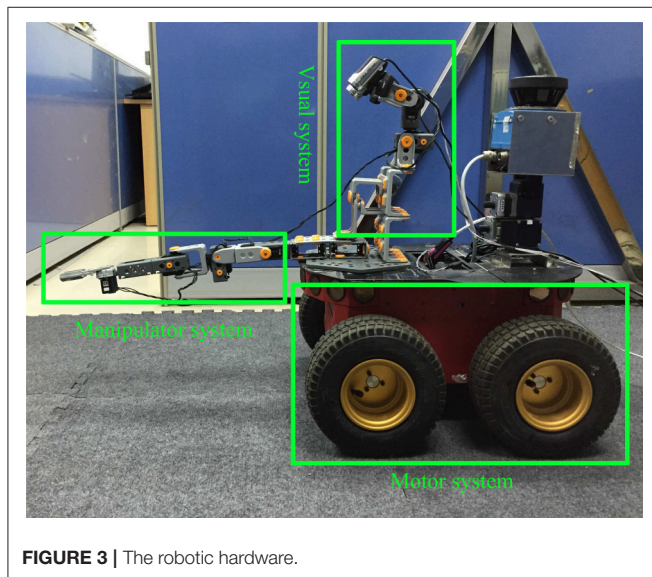
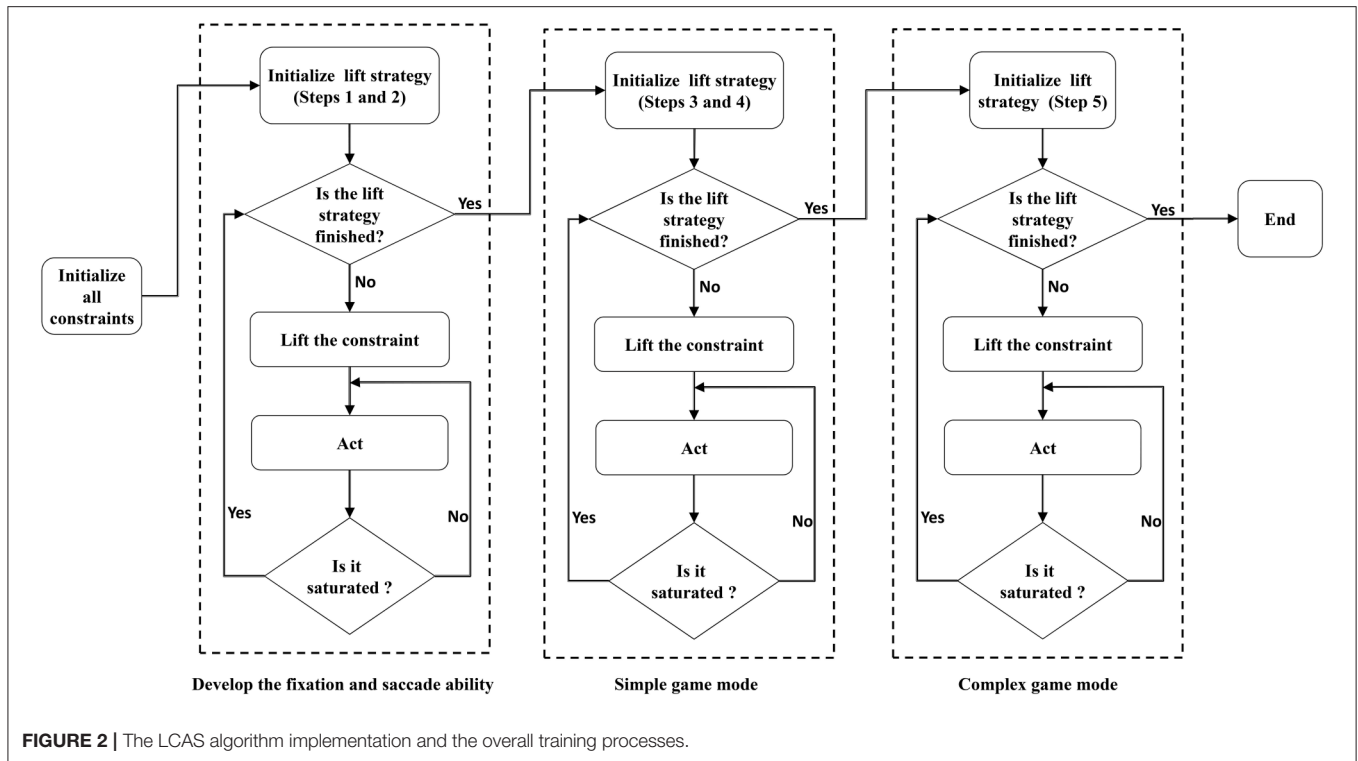
**TABLE 3 |** The robot's lift-constraint strategy.

Lift sequence	Developed ability
1. Visual resolution	Fixation ability
2. Eye joint	Saccade ability
3. Shoulder joint, Elbow joint	Hand-eye coordination
4. Wrist joint, gripper joint	Near-body grasping
5. Wheel joint	Mobile grasping

eye joint constraint is lifted. The robot learns saccade ability by eye joint movement. (3) Then, the arm joint constraint is lifted to allow for the movement and sensory abilities of the robot's arm. After the motor babbling stage, the robot builds hand-eye coordination. Accordingly, the robot executes the reaching action. (4) From this, the tactile sensor constraint in the arm is removed after the robot builds hand-eye coordination. Based on hand-eye coordination, the robot detects whether the object in the gripper uses the tactile sensor. At this stage, the robot learns near-body grasping. (5) Finally, the wheel joint constraint is removed, and the robot has mobile ability. Then, the robot learns mobile grasping by building the mapping between visual and mobile space.

After the first two steps of training, the robot developed fixation and saccade abilities, and thus the robot can play the simple game (Chao et al., 2016). In the simple game mode, the third and fourth steps of the lift-constraint strategy are executed, and the robot develops hand-eye coordination and near-body grasping. The fifth step of the strategy is executed in the complex game mode, where the robot develops mobile grasping ability. The entire procedure for model training and execution is illustrated in **Figure 2**.





### 3.3. Mobile Robot Hardware System

The mobile robot's intelligent system is mainly composed of three subsystems: the visual subsystem, the manipulator, and the motor, as demonstrated in **Figure 3**. The detailed functions of the three systems discussed below.

1. **The visual subsystem.** The robot's visual system performs two tasks including finding the object and locating the object. Firstly, the visual system analyzes the image color information

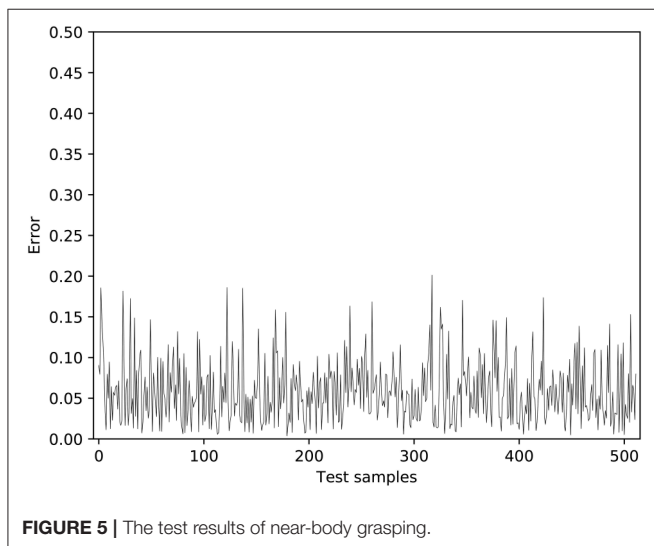
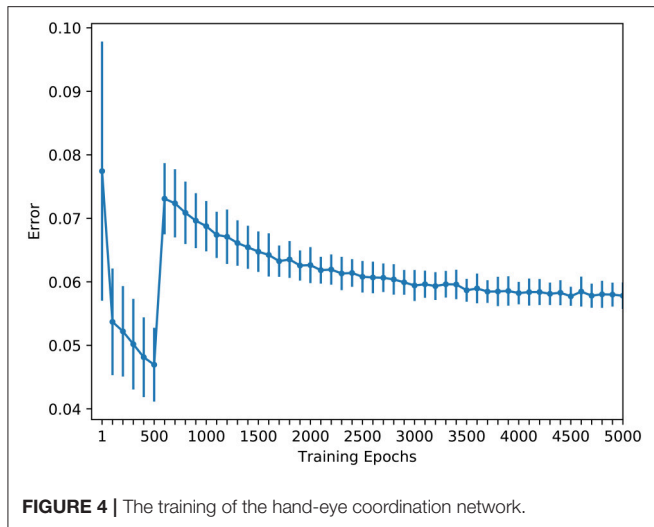
captured from the robot's two eyes, and detects whether the target is in the field of view. Secondly, if the target is in the field of view, the visual system, through fixation ability, acquires the retinal position of the target,  $S(x_1, y_1, x_2, y_2)$ , wherein  $s_l(x_1, y_1)$ , and  $s_r(x_2, y_2)$  express the coordinates of the target in the left and right eyes, respectively. The combination of the left and right eyes represents the target retinal position. Finally, the visual system, using the saccade ability, combines the retinal coordination,  $S(x_1, y_1, x_2, y_2)$ , and the eye joint,  $S_h(j_5, j_6)$ , to generate the visual space coordination,  $P(\gamma, \theta)$ .

2. **The manipulator subsystem.** The physical structure of the manipulator system simulates the human arm structure. It contains four joints, denoted as  $(j_1, j_2, j_3, j_4)$ . Its first three joints  $(j_1, j_2, j_3)$  correspond to the human arm joints (shoulder, elbow and wrist), respectively. These three joints construct the movement space of the robot arm. The  $j_4$  joint represents the gripper of the manipulator and is used to simulate the grasp ability of the human palm.
3. **The motor subsystem.** The motor system consists of four wheels, which enable the robot to execute mobile actions, such as forward motion, backward motion, and turns. Because only the front two wheels of the mobile robot have a motor, the robot's moving motor is denoted as  $M(m_1, m_2)$ . The robot controls its movements by changing the value of  $M(m_1, m_2)$ .

### 3.4. Game Processes

#### 3.4.1. Simple Game Process

The simple game mode requires the robot to pick up balls that are scattered within the work range of the manipulator. In this



game mode, the robot's eyes only focus on the work range of the manipulator. After it implements the first two steps of the lift-constraint strategy and develops fixation and saccade abilities, the robot can participate in the simple game mode, as shown in the center frame of **Figure 2**. When the robot plays the simple game, it must build a mapping from the visual space to the manipulator's movement space. This mapping is simulated by one of the two RBF networks,  $R_1$ . Different from the experiment hardware used by Caligiore et al. (2014), the robot is not a humanoid robot in this work. Therefore, it is impossible to use real data collected from infants, and instead the training samples are re-collected based on the wheeled robot. The training sample of this network consists of the visual space coordination,  $S_h(p, t)$ , and the manipulator's joint value,  $(j_1, j_2, j_3, j_4)$ . In the RBF network, the center positions of the RBF neurons are determined by a K-means algorithm, and the number of RBF neurons is set to twice the number of input dimensions. In this work, the Gaussian kernel is applied in the radial basis function. The calculation

of the RBF network is shown in Equations (5) and (6), where  $y(x)$  denotes the network's output joint value, and  $w_i$  denotes the weights of the hidden layer,  $\phi_i$  denotes a radial basis function,  $\delta$  denotes the width of the Gaussian kernel, and it is empirically set at  $\sqrt{5}$ .

$$y(x) = \sum_{i=1}^i w_i \phi_i \quad (2)$$

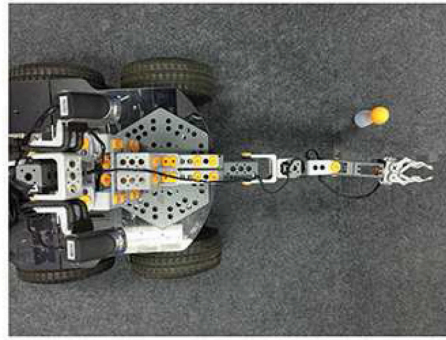
$$\phi_i(x) = \exp\left(-\frac{\|x - x_i\|^2}{2\delta^2}\right), \delta > 0 \quad (3)$$

In the process of training the robot's hand-eye coordination, a yellow ball is placed as a target object in the gripper of the manipulator. The ball moving randomly with the manipulator, and the sample  $e(P, M_a)$  are collected in each movement. The sample obtained is placed in the sample pool,  $E(e_1, e_2, e_3, \dots, e_n)$ . When the number of samples in the sample pool reaches a fixed number, some samples are randomly selected as the training samples for the  $R_1$  network. After that, the  $R_1$  network is trained using the Backpropagation algorithm. During the collection of the sample, the values of the wrist,  $j_3$ , and the gripper,  $j_4$ , are fixed, because the shoulder  $j_1$  and the elbow  $j_2$  already represent most of the movement space of the robot, while the wrist,  $j_3$ , is more involved in the grasping action (Marini et al., 2016). When the  $R_1$  network training is saturated, the robot develops a basic hand-eye coordination capability, and the wrist,  $j_3$ , and the gripper,  $j_4$ , can be released. After this constraint is removed, the samples in the sample pool must be re-collected. The  $R_1$  network uses these newly collected samples for further training, and, finally, the robot develops near-body grasping ability.

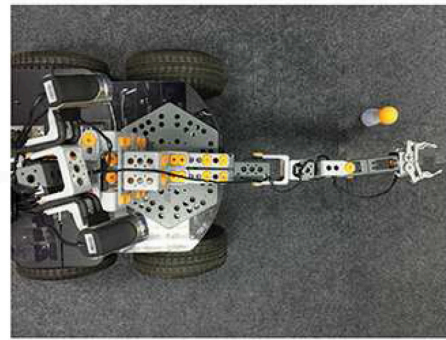
### 3.4.2. Complex Game Process

When the learning status is stable in the simple mode, the robot switches to the complex game mode using whole field view. The procedure for the complex mode is shown in the right-hand box of **Figure 2**. In the complex mode, the balls are scattered within the visual range of the robot, but not in the work range of the manipulator. To pick up these balls, the robot must relocate itself. The mapping relationship between the visual space and the robot's mobile space is built in the complex mode. This mapping is simulated by another RBF neural network,  $R_2$ . The training method and the set of parameters are the same as those used in the  $R_1$  network.

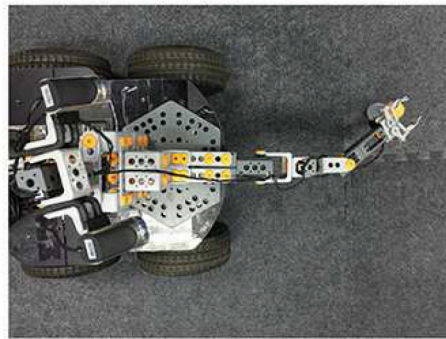
The training samples of the  $R_2$  network are based on two sequences: (1) the movement trajectory of the target in the robot's visual space,  $PS(p_{s1}, p_{s2}, \dots, p_{st})$ , and (2) the variation sequence of the robot's moving motor value  $MS(M_1, M_2, \dots, M_t)$ . In  $PS$ ,  $p_{st}$  denotes the coordinate distance of the ball in the robot's visual spaces when the robot moves from step  $t$  to step  $t + 1$ . The values of  $p_{st}$  is determined by Equation (4), where  $p_t(\gamma, \theta)$  denotes the position of the target in the visual space at step  $t$ , and  $p_{t+1}(\gamma, \theta)$  denotes the position at step  $t + 1$ . Likewise, the change of the motor value from step  $t$  to step  $t + 1$  is represented as  $M_t$ . So, the former  $n$ -step movement trajectory and the accumulated change of the motor are expressed by Equations (5) and (6) respectively, where  $PA_n$  denotes the accumulated distance from the target to



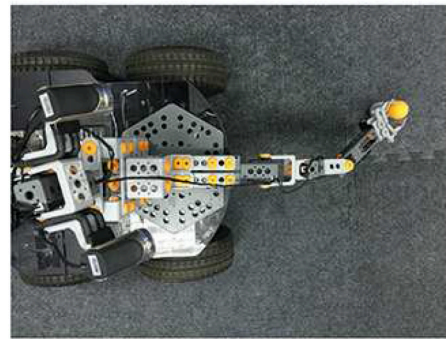
Step 1: Detect an object



Step 2: Saccade to the object



Step 3: Manipulator move towards the target



Step 4: Reach and grasp the object

**FIGURE 6** | The procedure in simple game mode.

the robot when the robot moves  $n$  steps, and  $MA_n$  denotes the accumulated change of the motor.

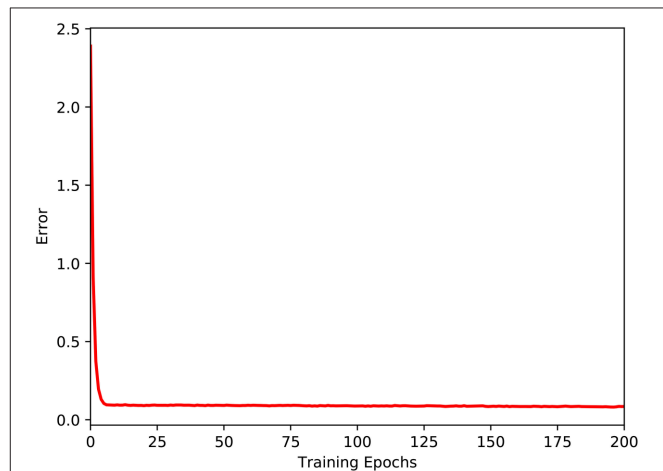
$$p_{st} = p_{t+1}(\gamma, \theta) - p_t(\gamma, \theta) \quad (4)$$

$$PA_n = P_{n+1} - P_1 \quad (5)$$

$$MA_n = \sum_{n=1}^n M_n, \quad n \leq D \quad (6)$$

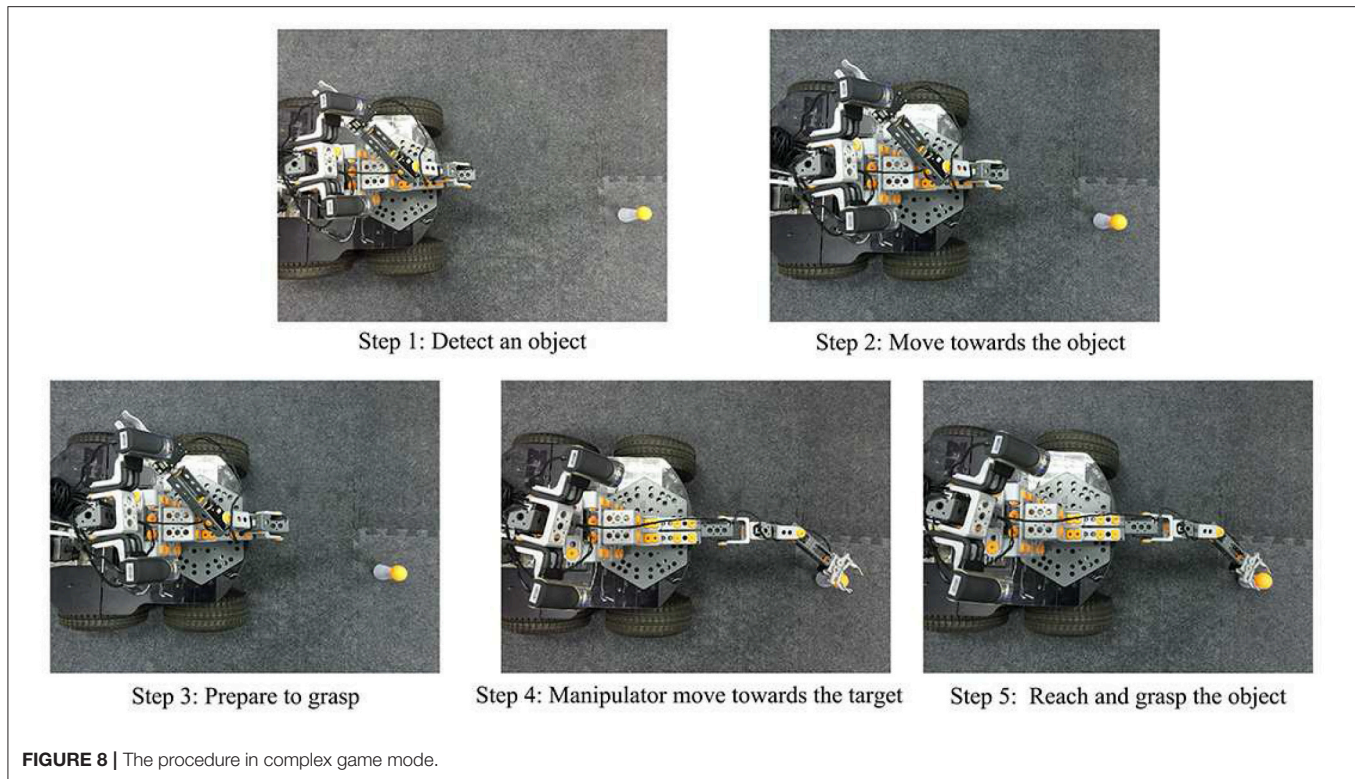
In the complex game, a target is placed within the field of the robot's vision, rather than within the manipulator's work field. Then, the robot is set to randomly move  $n$  steps. If the ball enters the manipulator work field occasionally during the  $n$  steps, the entire movement trajectory is chosen as a sample. However, if the ball is out of the field of robot's vision during the  $n$  steps, this trajectory is abandoned, and the target is randomly placed in a new position to start the iteration again.

This stage also requires the following two additional restrictions on the mobility of the robot. (1) Because the accuracy of the robot hardware is limited, the number of mobile steps  $n$  in a task must be less than the threshold, denoted as  $D$ . If the number of steps is too big, the accumulated error of the motor will be very large. (2) If the target disappears from the visual field during the robot moving, the task is considered to be a failure, and a new task is started by resetting the game. When the robot reaches the target position

**FIGURE 7** | The training of the eye-mobile coordination network.

within the number of steps,  $t$ , less than the threshold, the  $PA_n$  and the  $MA_n$  are combined into a sample  $e(PA_n, MA_n)$ . The training  $R_2$  network begins after enough samples have been collected in the sample pool  $E(e_1, e_2, e_3, \dots, e_n)$ . After the mapping relationship is properly established by training the  $R_2$  network, the robot develops mobile grasping ability. When





performing a mobile grasping task, the robot completes the task in two steps. Firstly, the robot moves toward the target within the work range of the manipulator using the trained  $R_2$  network. Then, the robot uses  $R_1$  to complete near-body grasping.

## 4. EXPERIMENTS AND ANALYSIS

### 4.1. The Simple Game

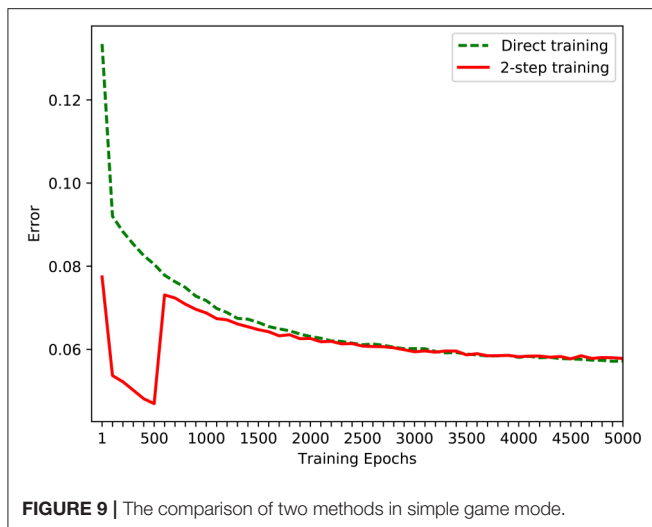
In order to train the hand-eye coordination network, 2,200 training samples and 512 test samples were collected. The error change during the training process is shown in **Figure 4**. In this figure, each circle denotes the average error after the experiment has been ran for 50 times; and each vertical bar denotes its standard deviation. The training error quickly reduced before the wrist and gripper joints were released. The average error was reduced to  $<0.05$  after just 500 training epochs, which is the point that the robot has successfully learned the hand-eye coordination. Then, the constraints of the wrist and gripper joints were released and the training of the robot's near-body grasping capability was initiated. **Figure 4** shows that, in the near-body grasping training stage, the training error begins to decline slowly after a rapid increase, eventually converging to about 0.06. After the network finished training, we tested it with 512 test samples, with the results shown in **Figure 5**. The overall average error for the 512 tests is about 0.07. Given that it is generally a success if the average error is  $<0.1$  in robot hardware systems, it is clear that the proposed robot has successfully learned the near-body grasping skill.

**Figure 6** shows the robot's performance during the game after it has learned the near-body grasping skill. In the first step, the robot detects whether the target is within the working range of the manipulator. If the target is within that range, the robot proceeds to the second step, where it views the target, by a saccade and obtains the exact position of the target within its field of vision. After that, the robot maps the visual position of the target into the movement space of the manipulator, and drives the manipulator toward the target position. Finally, the manipulator reaches and grasps the target.

### 4.2. The Complex Game

The eye-mobile network was trained after successfully trained the hand-eye network. In this stage, 600 samples were collected, of which 500 samples were used for training and the remaining 100 samples for testing. The threshold,  $D$ , is set as 7. This experiment has also been run for 50 times. Changes of the average error in the training process are shown in **Figure 7**. As seen in **Figure 7**, the training error immediately declines rapidly and reaches a stable minimum, indicating that the mapping between the robot's visual and mobile spaces is not very complicated.

**Figure 8** illustrates the robot's performance during the complex game mode after the eye-mobile network has finished training. In Step 1, the robot uses the visual system to obtain the position of target and detect whether the target lies within the working range of the manipulator. In Step 2, if the target is not in the working range of the manipulator, the robot drives the mobile system toward the target position until the target appears in the working range of the manipulator. In Step 3, after the target



appears in the working range of the manipulator, the robot stops moving and reacquires the visual position of the target. In Step 4, the robot drives the manipulator toward the target, and in Step 5, the robot reaches and grasps the target.

### 4.3. Performance Analysis and Comparative Study

The experiment discussed above allows the robot, step by step, to learn the near-body grasping skill using a developmental approach. In order to facilitate comparative study, another experiment with the same target of near-body grasping skill was designed using conventional direct training in the simple game mode. The performance of these two methods is compared in **Figure 9**, where the dotted line is the change of average training error that is learned directly in the simple game mode, and the solid line is the change of average error using the developmental method. As the results shown in **Figure 9**, in the first 2,000 epochs, the training efficiency of development method is higher than that of the direct training method. After 2,000 training epochs, the training efficiencies of these two methods achieve at a close error value. This phenomena proves that our approach is superior to the conventional method in learning efficiency. In the experiment using the developmental method, in the first 500 epochs, the network was trained using training samples for a robot whose wrist and gripper joints were constrained. After the error is less 0.05, the constraint was removed and the entire range of samples was used to retrain the network. As shown in **Figure 9**, as the number of epochs increases, the error rate for both the 2-step developmental and direct training approaches decreases. However, over the entire range of epochs, the error decreasing rate of the developmental approach is faster than that of the direct approach, indicating that using the developmental method in a game improves the learning efficiency of a robot.

To summarize, with the experiment of playing game, the robot successfully learned the mobile grasping ability by playing simple and complex games. Through the above experiments, we can conclude the following two results: (1) the proposed approach

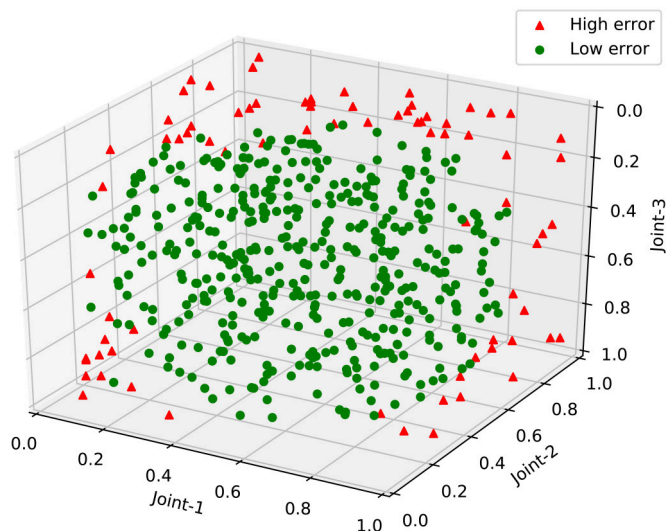
enables robots to learn skills by modeling the play activities during human infant development. (2) The developmental method with the game elements improves the robot's learning efficiency.

### 4.4. Comparison and Discussion

A comparison of **Figure 4** with **Figure 7** shows that the error rate for the eye-mobile network decreases more rapidly than that for the hand-eye network. However, in the early training epochs, the error rate is higher for the eye-mobile network than for the hand-eye network, because the mobile system has only two joints, but the manipulator has four. Therefore, mapping from visual space to mobile space is simpler than mapping to the manipulator movement space. On the other hand, fewer dimensions also make the output more sensitive to the input values. In **Figure 7**, the reason of the rapid error decreasing in  $R_2$  is that a simplified motor mode, mapping the visual space to robot's motor position space, is used in this work. The target's visual coordination and the robot's wheel movement trajectory are collected as training samples, in which, the motor mobile value has only two dimensions. The mapping between the robot's visual and mobile spaces is not complicated. However, if the work uses the mobile platform's dynamic control model, which can support the acceleration control for our robot, the network will require more learning time for the more complicated control. In **Figure 4** the error rate rises rapidly after the wrist and gripper joints are released at the five hundredth epoch, because the mapping becomes more complex. However, the error rate after the rapid increase is still lower than that for the directly learning approach, proving that the learning in the previous stage is helpful for the next learning stage.

After testing the robot's near-body grasping ability, we further analyze the test results from the perspective of the manipulator's movement space. Because the gripper joint has only two ways to open and close, it has little effect on the variation of the manipulator in the movement space. Therefore, we analyze just the first three joints of the manipulator. The analysis results are shown in **Figure 10**. Triangles represent the test sample for which the error rate is  $>0.1$ ; circles represent the others. **Figure 10** shows that most of the high error actions have at least two joints and an angle value near the extremum. In other words, these actions, distributed around the edge of the movement space, may be due to the instability of control when the manipulator's servos are near the maximum and minimum angles. Instead, the robot's grasping errors are generally below 0.1 at all other places. By removing the hardware factor, we assume that the robot has built the mapping from the visual space to the manipulator movement space.

For comparison, many developmental models used humanoid robots as experimental platforms (Marocco et al., 2010; Shaw et al., 2014; Morse and Cangelosi, 2017) in particular, several of them are infant-like robots. Those humanoid platforms directly benefit from important theories on developmental psychology, so that these platforms can easily reproduce several behavioral patterns or validate new hypotheses. In contrast, the shape and configuration of a mobile manipulator are very different from those of a human; therefore, the developmental theories need to



**FIGURE 10 |** The results of the near-body grasping test.

be adjusted to fit the robot system. However, currently, the mobile manipulators are competent to practical applications, which also requires the robots to have cognitive abilities. Thus, new developmental theories validated by mobile manipulators can be rapidly applied in real-life applications. Moreover, our work focus on using the “Play” strategy to create the mobile reaching ability for our robot, with the two game modes created. Our system not only involves spontaneous and intrinsically motivated exploration of actions and objects in varying contexts (Lee, 2011) but also, contains developmental characteristic by applying the “LCAS” developmental learning algorithm. Without setting specific goals, the robot uses its learning status to develop from the simple game mode to the complex one. The combination of developmental robotics and play modeling leads our robot to have faster learning rate.

## 5. CONCLUSION

Scientists have found that infants develop a number of skills when playing games in infant developmental research. In this paper, we combined these infant development theories with the LCAS algorithm to generate a developmental algorithm, which does not require specifically defined developmental goals for a robot. We designed two game modes and two RBF neural networks to simulate the procedures necessary for a robot to play in these game modes, with the support of a developmental strategy. The experiments demonstrated that a robot successfully learned moving and grasping skills. From results analysis and comparison, it can be concluded that: (1) a robot can successfully learn near-body grasping and moving grasping skills through play, and (2) in regard to a robot learning these skills, the developmental approach reduces the complexity and accelerates the learning speed.

Our model also has some limitations which can be mitigated in the future. For instance, in order to implement the hand-eye

coordination system rapidly, our model uses an open-loop method, which may lead to several failed grasping. In addition, our model does not use the data obtained from real infants. In order to address these, a close-loop method may be used to improve the success rate of grasping. In addition, as an infant develops skills through play, the infant's intrinsic motivation and ability to imitate are closely related to that play (Santucci et al., 2013; Oudeyer et al., 2016). In other words, infants achieve unsupervised learning in their environment through intrinsic motivation, which plays an important role in the control of the infancy learning stage transformation. However, infants learn new skills faster than other babies if they have a strong ability to imitate during the learning process. Therefore, the applicability of the proposed system can be extended in the future by incorporating intrinsic motivation and ability.

## AUTHOR CONTRIBUTIONS

RW performed the experiments and wrote the manuscript; ZZ and FC designed the robotic learning approach; CZ provided psychological analysis on the experimental data; CL designed the robotic control system; and LY analyzed the experimental results and edited the manuscript.

## ACKNOWLEDGMENTS

This work was supported by the Major State Basic Research Development Program of China (973 Program) (No. 2013CB329502), the Fundamental Research Funds for the Central Universities (No. 20720160126), the National Natural Science Foundation of China (No. 61673322 and 61673326), and Natural Science Foundation of Fujian Province of China (No. 2017J01129). The authors would like to thank the reviewers for their invaluable comments and suggestions, which greatly helped to improve the presentation of this paper.



## REFERENCES

- Adolph, K. E., and Joh, A. S. (2007). "Motor development: How infants get into the act," in *Introduction to Infant Development*, eds A. Slater and M. Lewis (New York, NY: Oxford University Press), 63–80.
- Asada, M., MacDorman, K. F., Ishiguro, H., and Kuniyoshi, Y. (2001). Cognitive developmental robotics as a new paradigm for the design of humanoid robots. *Rob. Auton. Syst.* 37, 185–193. doi: 10.1016/S0921-8890(01)00157-9
- Baillie, J.-C. (2016). "Artificial intelligence: The point of view of developmental robotics," in *Fundamental Issues of Artificial Intelligence*, ed V. C. Müller (Berlin: Springer), 413–422.
- Baldassarre, G., and Mirolli, M. (eds.). (2013). "Intrinsically motivated learning systems: an overview," in *Intrinsically Motivated Learning in Natural and Artificial Systems* (Berlin: Springer), 1–14.
- Berthouze, L., and Lungarella, M. (2004). Motor skill acquisition under environmental perturbations: on the necessity of alternate freezing and freeing of degrees of freedom. *Adapt. Behav.* 12, 47–64. doi: 10.1177/105971230401200104
- Caligiore, D., Mustile, M., Cipriani, D., Redgrave, P., Triesch, J., De Marsico, M., et al. (2015). Intrinsic motivations drive learning of eye movements: an experiment with human adults. *PLoS ONE* 10:e0118705. doi: 10.1371/journal.pone.0118705
- Caligiore, D., Parisi, D., and Baldassarre, G. (2014). Integrating reinforcement learning, equilibrium points, and minimum variance to understand the development of reaching: a computational model. *Psychol. Rev.* 121:389. doi: 10.1037/a0037016
- Cangelosi, A., Schlesinger, M., and Smith, L. B. (2015). *Developmental Robotics: From Babies to Robots*. Cambridge, MA: MIT Press.
- Chao, F., Lee, M. H., Jiang, M., and Zhou, C. (2014a). An infant development-inspired approach to robot hand-eye coordination. *Int. J. Adv. Rob. Syst.* 11:15. doi: 10.5772/57555
- Chao, F., Lee, M. H., and Lee, J. J. (2010). A developmental algorithm for ocular-motor coordination. *Rob. Auton. Syst.* 58, 239–248. doi: 10.1016/j.robot.2009.08.002
- Chao, F., Wang, Z., Shang, C., Meng, Q., Jiang, M., Zhou, C., et al. (2014b). A developmental approach to robotic pointing via human-robot interaction. *Inf. Sci.* 283, 288–303. doi: 10.1016/j.ins.2014.03.104
- Chao, F., Zhang, X., Lin, H.-X., Zhou, C.-L., and Jiang, M. (2013). Learning robotic hand-eye coordination through a developmental constraint driven approach. *Int. J. Autom. Comput.* 10, 414–424. doi: 10.1007/s11633-013-0738-5
- Chao, F., Zhu, Z., Lin, C.-M., Hu, H., Yang, L., Shang, C., et al. (2016). Enhanced robotic hand-eye coordination inspired from human-like behavioral patterns. *IEEE Trans. Cogn. Dev. Syst.* doi: 10.1109/TCDS.2016.2620156. [Epub ahead of print].
- Earland, K., Lee, M., Shaw, P., and Law, J. (2014). Overlapping structures in sensory-motor mappings. *PLoS ONE* 9:e84240. doi: 10.1371/journal.pone.0084240
- Galbraith, B. V., Guenther, F. H., and Versace, M. (2015). A neural network-based exploratory learning and motor planning system for co-robots. *Front. Neurobot.* 9:7. doi: 10.3389/fnbot.2015.00007
- Gogate, L. (2016). Development of early multisensory perception and communication: from environmental and behavioral to neural signatures. *Dev. Neuropsychol.* 41, 269–272. doi: 10.1080/87565641.2017.1279429
- Grinke, E., Tetzlaff, C., Wörgötter, F., and Manoonpong, P. (2015). Synaptic plasticity in a recurrent neural network for versatile and adaptive behaviors of a walking robot. *Front. Neurobot.* 9:11. doi: 10.3389/fnbot.2015.00011
- Guerin, F., Kruger, N., and Kraft, D. (2013). A survey of the ontogeny of tool use: from sensorimotor experience to planning. *IEEE Trans. Auton. Ment. Dev.* 5, 18–45. doi: 10.1109/TAMD.2012.2209879
- Hart, S., and Grupen, R. (2011). Learning generalizable control programs. *IEEE Trans. Auton. Ment. Dev.* 3, 216–231. doi: 10.1109/TAMD.2010.2103311
- Hülse, M., McBride, S., Law, J., and Lee, M. (2010). Integration of active vision and reaching from a developmental robotics perspective. *IEEE Trans. Auton. Ment. Dev.* 2, 355–367. doi: 10.1109/TAMD.2010.2081667
- Jiang, L., Meng, D., Yu, S.-I., Lan, Z., Shan, S., and Hauptmann, A. (2014). "Self-paced learning with diversity," in *Advances in Neural Information Processing Systems* 27, eds Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger (Montreal, QC: Curran Associates, Inc), 2078–2086.
- Law, J., Lee, M., and Hülse, M. (2010). "Infant development sequences for shaping sensorimotor learning in humanoid robots," in *10th International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems* (Glumslöv), 65–72.
- Law, J., Lee, M., Hülse, M., and Tomassetti, A. (2011). The infant development timeline and its application to robot shaping. *Adapt. Behav.* 19, 335–358. doi: 10.1177/1059712311419380
- Law, J., Shaw, P., Earland, K., Sheldon, M., and Lee, M. (2014a). A psychology based approach for longitudinal development in cognitive robotics. *Front. Neurobot.* 8:1. doi: 10.3389/fnbot.2014.00001
- Law, J., Shaw, P., and Lee, M. (2013). A biologically constrained architecture for developmental learning of eye-head gaze control on a humanoid robot. *Auton. Rob.* 35, 77–92. doi: 10.1007/s10514-013-9335-2
- Law, J., Shaw, P., Lee, M., and Sheldon, M. (2014b). From saccades to grasping: a model of coordinated reaching through simulated development on a humanoid robot. *IEEE Trans. Auton. Ment. Dev.* 6, 93–109. doi: 10.1109/TAMD.2014.2301934
- Lee, M., Law, J., and Hülse, M. (2013). "A developmental framework for cumulative learning robots," in *Computational and Robotic Models of the Hierarchical Organization of Behavior*, eds G. Baldassarre and M. Mirolli (Berlin: Springer), 177–212.
- Lee, M. H. (2011). "Intrinsic activity: from motor babbling to play," in *2011 IEEE International Conference on Development and Learning (ICDL)* (Frankfurt), Vol. 2, 1–6.
- Lee, M. H., Meng, Q., and Chao, F. (2007). Staged competence learning in developmental robotics. *Adapt. Behav.* 15, 241–255. doi: 10.1177/1059712307082085
- Lungarella, M., Metta, G., Pfeifer, R., and Sandini, G. (2003). Developmental robotics: a survey. *Connect. Sci.* 15, 151–190. doi: 10.1080/09540090310001655110
- Marini, F., Squeri, V., Morasso, P., and Masia, L. (2016). Wrist proprioception: amplitude or position coding? *Front. Neurobot.* 10:13. doi: 10.3389/fnbot.2016.00013
- Marocco, D., Cangelosi, A., Fischer, K., and Belpaeme, T. (2010). Grounding action in the sensorimotor interaction with the world: experiments with a simulated iCub humanoid robot. *Front. Neurobot.* 4:7. doi: 10.3389/fnbot.2010.00007
- Morse, A. F., and Cangelosi, A. (2017). Why are there developmental stages in language learning? A developmental robotics model of language development. *Cogn. Sci.* 41, 32–51. doi: 10.1111/cogs.12390
- Oudeyer, P.-Y. (2017). What do we learn about development from baby robots? *Wiley Interdiscipl. Rev. Cogn. Sci.* 8:e1395. doi: 10.1002/wcs.1395
- Oudeyer, P.-Y., Gottlieb, J., and Lopes, M. (2016). Intrinsic motivation, curiosity, and learning: Theory and applications in educational technologies. *Prog. Brain Res.* 229, 257–284. doi: 10.1016/bs.pbr.2016.05.005
- Oudeyer, P.-Y., Kaplan, F., and Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE Trans. Evol. Comput.* 11, 265–286. doi: 10.1109/TEVC.2006.890271
- Salgado, R., Prieto, A., Bellas, F., and Duro, R. J. (2016). "Improving extrinsically motivated developmental robots through intrinsic motivations," in *2016 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*. (Cergy-Pontoise: IEEE), 154–155.
- Santucci, V. G., Baldassarre, G., and Mirolli, M. (2013). Which is the best intrinsic motivation signal for learning multiple skills? *Front. Neurobot.* 7:22. doi: 10.3389/fnbot.2013.00022
- Savastano, P., and Nolfi, S. (2013). A robotic model of reaching and grasping development. *IEEE Trans. Auton. Ment. Dev.* 5, 326–336. doi: 10.1109/TAMD.2013.2264321
- Shaw, P., Law, J., and Lee, M. (2014). A comparison of learning strategies for biologically constrained development of gaze control on an iCub robot. *Auton. Rob.* 37, 97–110. doi: 10.1007/s10514-013-9378-4
- Stoytchev, A. (2009). Some basic principles of developmental robotics. *IEEE Trans. Auton. Ment. Dev.* 1, 122–130. doi: 10.1109/TAMD.2009.2029989
- Wang, Z., Chao, F., Lin, H., Jiang, M., and Zhou, C. (2013). "A human-like learning approach to developmental robotic reaching," in *2013 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. (Shenzhen: IEEE), 581–586.
- Wang, Z., Xu, G., and Chao, F. (2014). "Integration of brain-like neural network and infancy behaviors for robotic pointing," in *2014 International Conference*



- on *Information Science, Electronics and Electrical Engineering (ISEEE)*, Vol. 3. (Sapporo: IEEE), 1613–1618.
- Weng, J. (2004). Developmental robotics: theory and experiments. *Int. J. Humanoid Robot.* 1, 199–236. doi: 10.1142/S0219843604000149
- Yan, W., Weber, C., and Wermter, S. (2013). Learning indoor robot navigation using visual and sensorimotor map information. *Front. Neurobot.* 7:15. doi: 10.3389/fnbot.2013.00015
- Yang, B.-H., and Asada, H. (1996). Progressive learning and its application to robot impedance learning. *IEEE Trans. Neural. Netw.* 7, 941–952. doi: 10.1109/72.508937
- Zhu, Z., Chao, F., Zhang, X., Jiang, M., and Zhou, C. (2015). “A developmental approach to mobile robotic reaching,” in *International Conference*

on *Intelligent Robotics and Applications*, (Portsmouth, UK: Springer), 284–294.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Wu, Zhou, Chao, Zhu, Lin and Yang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.